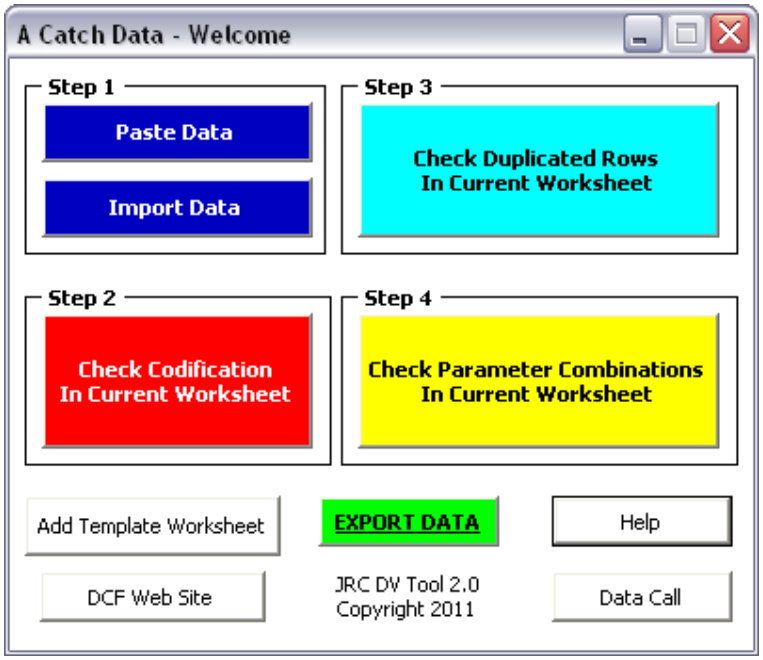
JRC Scientific and Technical Reports

# JRC Data Validation Tool for the Fisheries Effort Data Call through the EU Data Collection Regulation Framework (DCF)

DV Tool 2.0

User Manual

**Nikolaos Mitrakis**




EUR 24845 EN - 2011

JRC
EUROPEAN COMMISSION

ipsc
Institute for the Protection and Security of the Citizen

The mission of the JRC-IPSC is to provide research results and to support EU policy-makers in their effort towards global security and towards protection of European citizens from accidents, deliberate attacks, fraud and illegal actions against EU policies.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

**Contact information**
Address: Via Fermi 2749, T.P. 051 – Ispra (VA), 21027 - Italy
E-mail: nikolaos.mitrakis@jrc.ec.europa.eu
Tel.: 0332 786479
Fax: 0332 789658

http://ipsc.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/



JRC 64407

EUR 24845 EN
ISBN 978-92-79-20442-5
ISSN 1831-9424 (online)
ISSN 1018-5593 (print)
doi:10.2788/27561

Luxembourg: Publications Office of the European Union



*Printed in Italy*

## TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# 1. Introduction

The Data Validation (DV) tool is a set of macros developed in Visual Basic for Applications (VBA) and embedded in specifically designed template Excel Workbooks for the effort data call. The main purpose of this tool is to facilitate and support the Member States in uploading data which meet the requirements defined by DG Mare in the official DCF data call for fishing effort regime evaluations by STECF (Council Regulation 199/2008). The use of these Excel Template files is not mandatory. However, the data validations checks performed by the DV tool can highly reduce the number of erroneous records the file to be uploaded to the DCF web site may contain, and hence facilitate the uploading procedure.

The tool is capable of checking national data stored in Excel rows against certain codifications and rules as requested in the effort data call. The data can be part or the whole of the data to be uploaded to the DCF web site. The checks are for syntactic but also semantic errors. The majority of the checks concerns the use of valid codes referred to the various Appendices of the data call and the type of the data entered (numeric or text). Erroneous data are identified and can be easily corrected. However, the most important feature of the tool is the ability to check about the use of valid combinations of certain field values in accordance with specific fishery regulations.

Currently, there are four (4) template files available, named **catch.xlsm**, **effort.xlsm**, **speffort.xlsm** and **capacity.xlsm** for Office 2007 users, and four (4) template files, named **catch.xls**, **effort.xls, speffort.xls** and **capacity.xls** for Office 2003 users. These four files correspond to the four data tables as described in the effort data call. The tables requested by the data call are **A Catch**, **B Effort, C Spatial Effort** and **D Capacity**.

The current version of the DV tool is 2.0. However, the tool is still under development and the author seeks the feedback of the users in order to improve the tool according to their needs. Any changes in the data call imply coherent adaptation of the modules. The purpose of this user manual is to provide the concept of this tool and guidelines for its use. Since it is a tool integrated in Excel Workbooks, all the Excel functions are still available. The template files are available for download from the Data Collection Framework web site https://datacollection.jrc.ec.europa.eu/web/dcf/dc/effort.

## 1.1. Map of the document

The manual consists of the following sections. *Chapter 1* presents the purpose and the structure of the document. *Chapter 2* describes the DV tool (**catch.xlsm**) for the check of the records corresponding to the A Catch table of the data call, *Chapter 3* describes the tool for the B Effort table (**effort.xlsm**), *Chapter 4* the tool for the C Spatial Effort table (**speffort.xlsm**) and *Chapter 5*, the tool for the D Capacity table (**capacity.xlsm**). In *Chapter 2*, a detailed description of the four basic windows, which are common through all the files, is given:

- Welcome Screen
- Check Codification
- Check Duplications
- Check Combinations

The specific checks applied to the different types of data are described in each chapter. *Chapter 6* describes the procedure for correcting the errors encountered whereas *Chapter 7* describes how the users can export the data in the appropriate format. Finally, *Chapter 8* presents hints for Office 2003 users and *Chapter 9* includes some important notes for the users.

# 2. A Catch Data

When opening the catch.xlms template file, a **Security Warning** may appear with the message:

"Macros have been disabled. Options…"

This is a standard security option in all Microsoft's applications, since macros potentially can have access to your data. In order to continue with the use of the tool, the user should choose: "Options→Enable this content". It is advised not to change the default option not to enable macros. Although DV tool is a trusted application, other macros can potentially cause several problems. To change the security settings go to Developer→Code→Macro Security→Disable all macros with notification.

## 2.1. Welcome Screen

The **Welcome Screen** is the first and the main window that pops up when opening the Workbook (**Figure 1**). From here, the users can have access to all the check functions that are provided from the DV Tool.

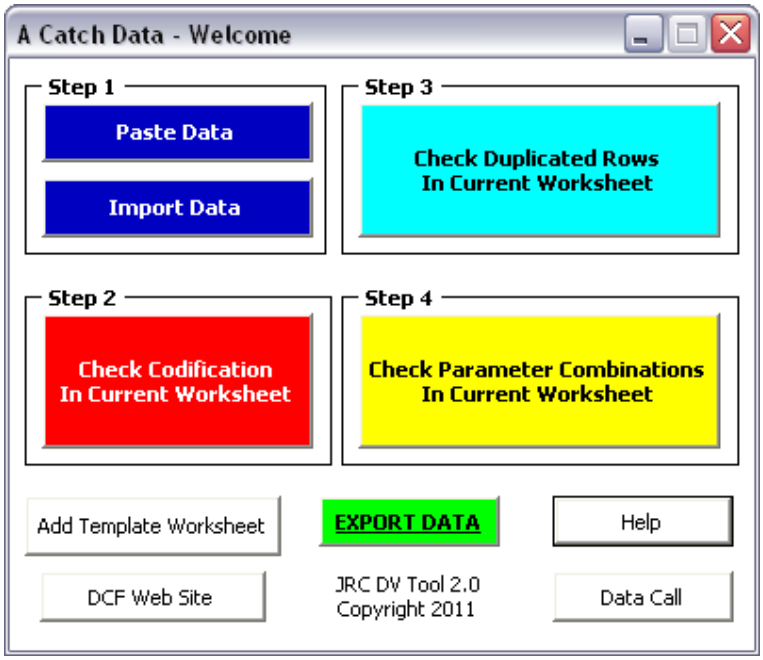



**Figure 1:** Welcome Screen of the DV tool for the Catch Data.

There are two main areas in this window. The first is consisting of a four step procedure, and the second is a support area. The users are strongly advised to follow each step of the procedure in order to guarantee the appropriate check of the data. The available buttons are:

- **Step 1**
  - → Paste Data
  - → Import Data
- **Step 2**
  - → Check Codification In Current Worksheet
- **Step 3**
  - → Check Duplicated Rows In Current Worksheet

- **Step 4**

  → Check Parameter Combinations In Current Worksheet

- **Export Data** (Exports the data in the Workbook to files ready to be uploaded to the DCF web site)
- **Add Template Worksheet** (Adds an accordingly formatted Worksheet in the Workbook)
- **DCF Web Site** (Click here to connect to the DCF web site)
- **Help** (Click here to download the User Manual)
- **Data Call** (Click here to download the effort data call official letter)

## 2.2. Step 1

In order to check if the data meet the requirements of the effort data call, the first step is to populate the Worksheet with the records of your data. The template Workbook by default has one template Worksheet, named **EFF_01_CATCH**, with fields as defined in the effort data call for the Catch table. Each of the columns is formatted according to the type of data the field describes. The user has two options for populating the rows of this Worksheet: **Paste** and **Import** Data.

**Paste Data**

In most of the cases, users keep their data or can easily export them in Excel files. To paste data from another Excel Worksheet, just open the Worksheet with the original data, select the rows with the data you want to check, select the template Worksheet, and click the blue **Paste Data** button. Regardless of which cell is selected in the template Worksheet, when clicking the **Paste Data** button, the data are pasted in the A2 cell, right down of the title row. In case user's data contains a header row, when pasted, the header row will be under the template Worksheet's header row. The user should delete it to avoid unnecessary errors.

**Import Data**

The second option is to import text files, tab, comma, semicolon or space delimited. The valid file types are *txt, *.csv and *.prn. When clicking the **Import Data** button, the default Dialog Window pops up where the user can chose which file to import. After choosing the file and clicking the OK button, the application asks for the delimited character (**Figure 2**). Enter according to **Table 1**.

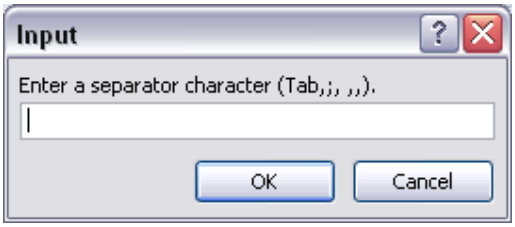



**Figure 2**: Input dialog box for separator character.

**Table 1**: Delimited characters for importing text files

| Input | Delimited Character |
|---|---|
| Tab | Tab |
| ; | Semicolon |
| , | Comma |
| | Space |

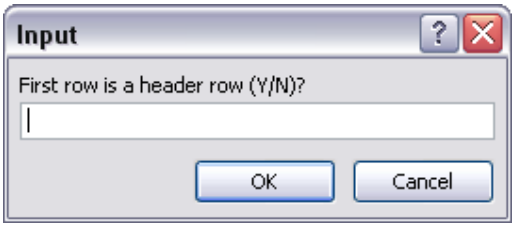



**Figure 3**: Input dialog box for header row.

The second input window (**Figure 3**) asks the user to reply if the first row of the data contains column headers. Enter **Y**(Yes) or **N**(No) accordingly. The application creates a new template Worksheet, with a header row, and fills the Excel cells with the data stored in the text file. If the text file contains a header row and the user wrongly chooses **Y** in the input dialog box, then the header line is inserted too. In this case, the line will be repeated under the template Worksheet's header row. The user should delete one header row to avoid unnecessary errors. The name of the new Worksheet is again **EFF_01_CATCH** with a space character following the name.

**Important**: A second option to import data as an Excel Worksheet is to use Excel's embedded import tool. The procedure to follow is to click on the **Add Template Worksheet**, choose from Excel's Data ribbon **Data → Get External Data,** choose all the columns and rows to import, choose the type of the data in the Columns (not General but Text or Numeric), and choose to import in the A2 Cell of the current Worksheet.

## 2.3. Step 2

After filling the cells of a template Worksheet with national data, the next step is to check the national data against the codification required by the effort data call.

*Check Codification In Current Worksheet*

When clicking the red button, the following window appears (**Figure 4**), and starts the procedure of checking if the data follows the codification scheme described in the effort data call.
The variables checked by this function are (case insensitive):

| | |
|---|---|
| COUNTRY | The valid codes are given in Appendix 1 of the data call. |
| YEAR | The entry should be a year, not after 2010 or earlier than 2003. |
| QUARTER | The quarter should be an integer of 1, 2, 3 or 4. |
| VESSEL_LENGTH | The valid codes are given in Appendix 2 of the data call. |
| GEAR | The valid codes are given in Appendix 3 of the data call. |
| MESH_SIZE_RANGE | The valid codes are given in Appendix 4 of the data call. |
| AREA | The valid codes are given in Appendix 5 of the data call. |
| SPECON | The valid codes are given in Appendix 6 of the data call. |
| SPECIES | The valid codes are given in Appendix 7 of the data call. |
| LANDINGS | The entries should be numerical values, not text. |
| DISCARDS | The entries should be numerical values, not text. |

The *Check Codification* window shows the number of records processed so far, together with a progress bar, which indicates the status of the Step 2 procedure. The time needed to check the codes is approximately 1 sec/1000 records on an Intel Core 2 TP5550 @ 1.66GHz CPU with 2 GB of RAM. However, the time also depends on the applications running on user's PC. The elapsed time is given in hours, minutes and seconds (hh:mm:ss format).

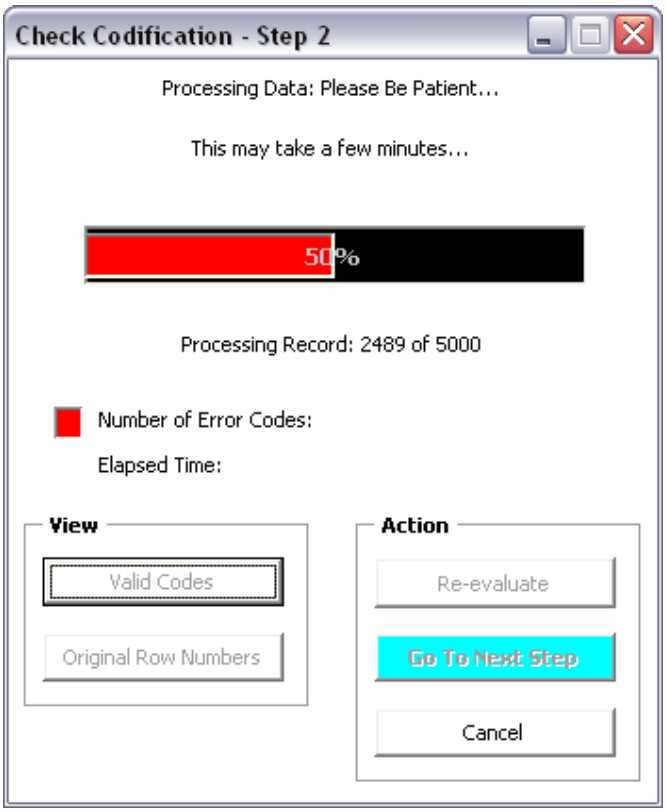



**Figure 4**: Check Codification window when checking for valid codes.

During the check, the only enabled button is **Cancel**. Clicking on this button while running the checks, forces the process to stop, and only the first rows will be checked.

After finishing this process, or clicking **Cancel**, the rest of the buttons are enabled. There are two different button areas, **View** and **Action**.

*View*

**Valid Codes** — A Worksheet is appearing, showing all the valid codes for your data. It is impossible to alter the entries of this Worksheet. The button label is changed to **Go Back - Hide Codes**. Clicking again on the button will transfer back in the initial Worksheet.

**Original Row Numbers** — Clicking this button will insert a new column (column A) on the left of the Worksheet, which holds the position of each record in the original file the user imported or pasted data from. This number allows verification against the initial data. The button label is turned into **Hide Original Row Numbers**. Clicking again on the button will hide the column from the Worksheet.

In order to correct the errors that may appear, go to *Chapter 6*, **Correcting the Errors.**

*Action*

**Re-evaluate** — Clicking on **Re-evaluate** always repeat the data checking.

**Go To Next Step** — If error free, the user will automatically proceed to the third step. In case several errors are detected, a message will appear warning that the errors should be corrected before proceeding to the next step.

**Cancel** Enabled during the check procedure. Clicking on it will force the procedure to stop.

In case the user successfully checks the data against the codification of the data call, the window given in **Figure 5** appears.

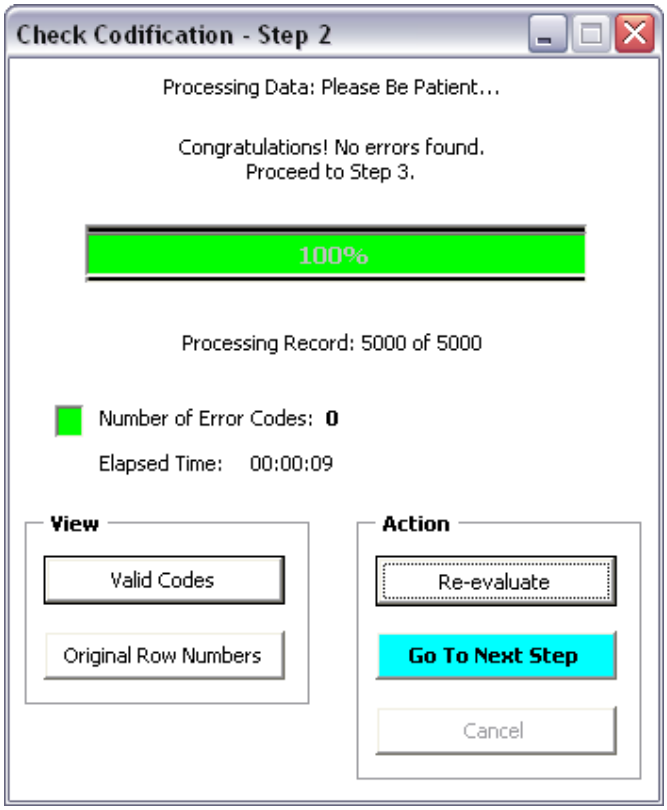



**Figure 5**: Successful check of the codification scheme followed.

## 2.4. Step 3

Click on *Check Duplicated Records In The Current Worksheet* on the *Welcome Screen*, or in **Go To Next Step** button in the *Check Codification* window to proceed with the examination of possible duplicated records in the aggregation level in the current Worksheet.

All duplicated records are coloured cyan (the cells that define the aggregation level) and pushed to the top. For the A Catch table the aggregation level is defined by the following parameters:

COUNTRY – YEAR – QUARTER – VESSEL_LENGTH – GEAR – MESH_SIZE_RANGE – AREA – SPECON – SPECIES

The time needed to check the presence of duplicated records is approximately 2 secs/1000 records on an Intel Core 2 TP5550 @ 1.66GHz CPU with 2 GB of RAM. The window is given in **Figure 6**.

The window has the same form as previous. All the buttons perform the same action as in the previous step. In case the user successfully checks the data for duplications, the window given in **Figure 7** appears.

## 2.5. Step 4

Click on *Check Parameter Combinations In The Current Worksheet* on the *Welcome Screen*, or in **Go To Next Step** button in the *Check Duplications* window to proceed with the examination of the combination of the following variables: **Gear** (GEAR), **Mesh size range** (MESH_SIZE_RANGE), **Area**

(AREA), **Specific condition** (SPECON). This is an important step, since these variables must have an appropriate combination of entries consistent with various fishery regulations. The time needed to check the parameter combinations is approximately 30 secs/1000 records on an Intel Core 2 TP5550 @ 1.66GHz CPU with 2 GB of RAM. The window is given in **Figure 8**.

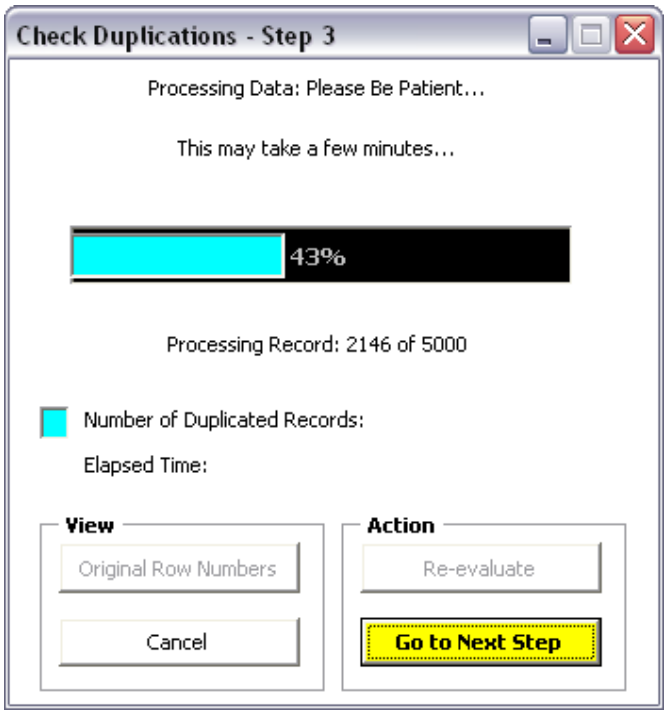



**Figure 6**: Check Duplications window.

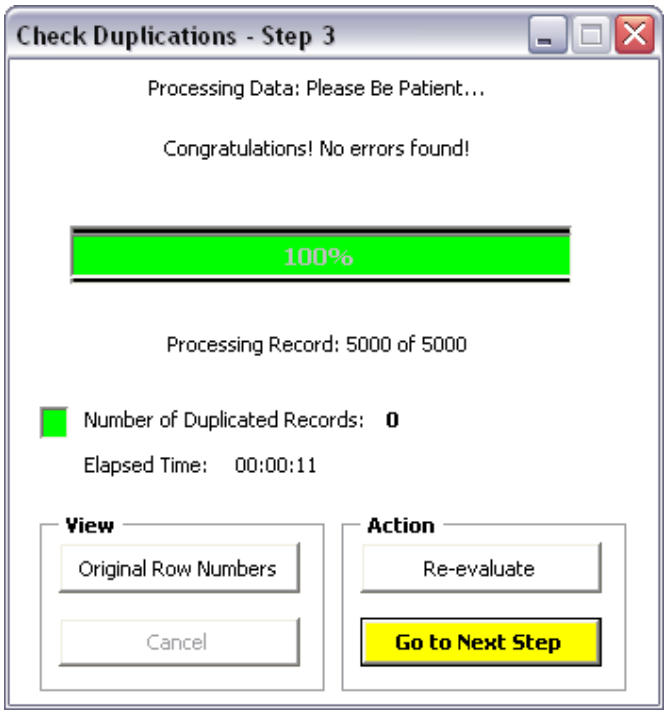



**Figure 7**: Successful check for duplicated records in the aggregation level.

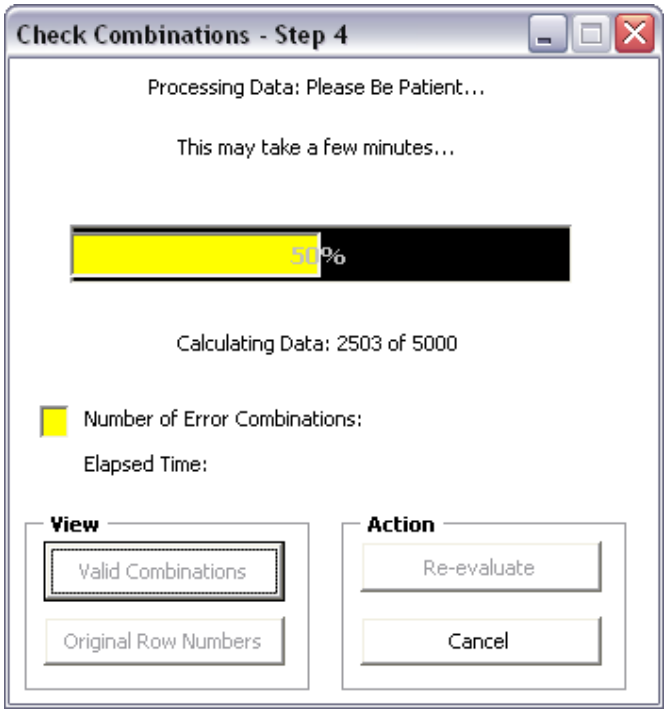



**Figure 8**: Check Combinations window when checking for valid parameter combinations.

The window has the same form as previous. All the buttons perform the same action as in the previous step. The difference lies in the **Valid Combinations** view.

**Valid Combinations** A Worksheet is appearing, showing all the valid combinations of the national data. You cannot alter the entries of this Worksheet. The user has the option to filter the data accordingly and check which the valid combinations are for certain variables entries. The button label is changed to **Go Back to Data**. Click again on the button to go back in the initial Worksheet.

If the procedure ends successfully and no error combinations are detected, the message given in **Figure 9** appears. Click on the "Yes" button to transfer to the DCF web site and upload the file, or click "No" to return to the *Welcome Screen.*

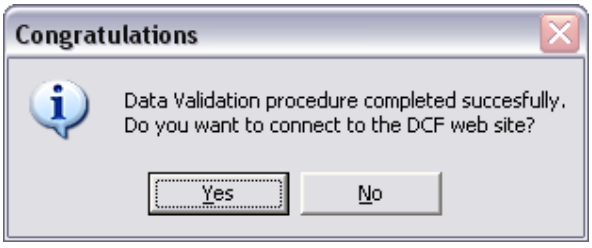



**Figure 9**: Successful termination of the data validation procedure.

To exit the *Welcome Screen* just click on the red X button in the right upper corner of the window. When exiting the message given in **Figure 10** appears. The user is reminded to delete all the temporary or unnecessary Worksheets in the Workbook, and keep only those having the corrected records of data.

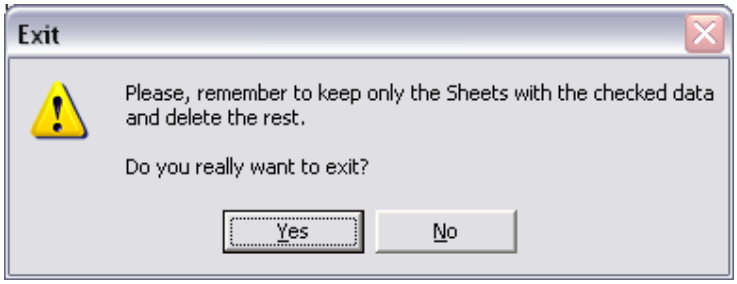



**Figure 10**: Exit DV Tool message.

To re-open the DV Tool when closed, just click on the button named "JRC DV Tool" on the "Home" ribbon, as showed in **Figure 11**.


**Figure 11**: The DV Tool button in Excel's Home ribbon.

## 2.6. Messages

Several messages can appear when using the DV Tool, which have as main purpose to warn or inform the user about the status of the procedure.

*Going directly to Steps 3 or 4 from the Welcome Screen*

The user has the option to skip Steps 2 and/or 3, the check for valid codes and duplicated errors, and move directly to Steps 3 and/or 4. This can be the case if for example, the user has already checked for valid codes and/or duplicate records, made all the necessary changes, and then closed and saved the Workbook. Next time that the user re-opens the Workbook will not have to check again for valid codes and/or duplications, since the records are considered to be correct as far as codification is concerned and/or duplications. Proceeding directly to Step 3 for example, forces the appearance of the message given in **Figure 12**.

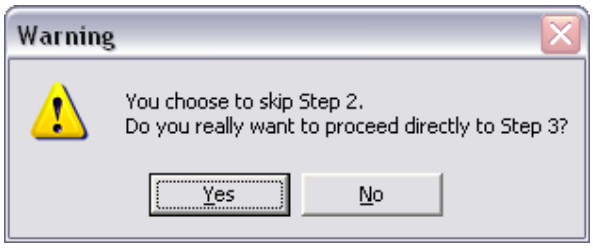



**Figure 12**: Warning message when clicking directly from the Welcome Screen on the Step 3 button.

In case the software detects some errors in Step 2 and the user chooses to proceed to Step 4 by clicking on *Check Parameter Combinations In The Current Worksheet,* the message shown in **Figure 13** will appear. If the user proceeds to Step 4 by clicking **OK** and Step 4 finishes successfully, then the message in **Figure 14** will appear, informing the user that there are still errors in some of the codes used in the records.

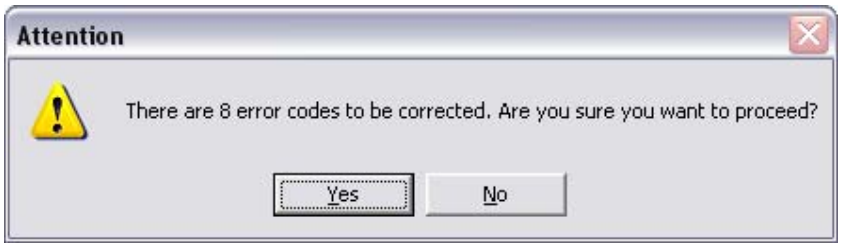



**Figure 13**: Warning message when the user proceeds to Step 4 when there are errors in Step 2.

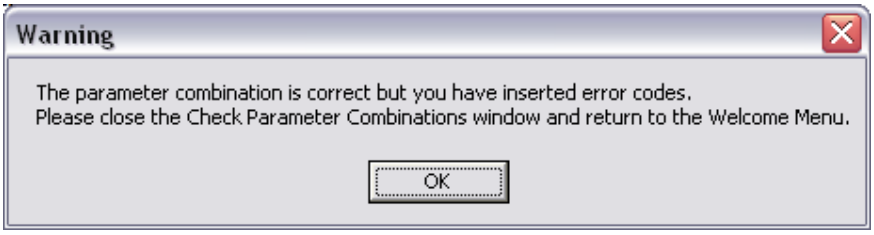



**Figure 14**: Warning message when Step 4 is completed successfully but there are still errors in Step 2.

When the program tries to open a file hosted in DCF's web site (such as the Data Call), the standard Microsoft's security message may appear (**Figure 15**) informing about possible risks of this action. The user should choose **OK**.


**Figure 15**: Microsoft's message when trying to open a file hosted in the DCF web site.

# 3. B Effort Data

## 3.1. Welcome Screen

The Welcome Screen is the same as in the case of the Catch Data table as shown in **Figure 16**. The window again consists of two main areas: a four step procedure, and a support area.

## 3.2. Step 1

For **Step 1**, please consider 2.2. The name of the default Worksheet is **EFF_02_EFF**.

## 3.3. Step 2

When clicking the red button, the *Check Codification* window appears, and starts the procedure of checking if the data follows the codification scheme described in the effort data call.

The variables checked by this function are (case insensitive):

| | |
|---|---|
| COUNTRY | The valid codes are given in Appendix 1 of the data call. |
| YEAR | The entry should be a year, not after 2010 or earlier than 2000. |
| QUARTER | The quarter should be an integer of 1, 2, 3 or 4. |
| VESSEL_LENGTH | The valid codes are given in Appendix 2 of the data call. |
| GEAR | The valid codes are given in Appendix 3 of the data call. |
| MESH_SIZE_RANGE | The valid codes are given in Appendix 4 of the data call. |
| AREA | The valid codes are given in Appendix 5 of the data call. |
| SPECON | The valid codes are given in Appendix 6 of the data call. |
| ALLOWED_ACTIVITY | The entries should be numerical values or null. Not text. |
| FISHING_ACTIVITY | The entries should be numerical values or null. Not text. |
| FISHING_CAPACITY | The entries should be numerical values or null. Not text. |
| NOMINAL_EFFORT | The entries should be numerical values or null. Not text. |
| GT_DAYS_AT_SEA | The entries should be numerical values or null. Not text. |
| NO_VESSELS | The entries should be numerical values or null. Not text. |

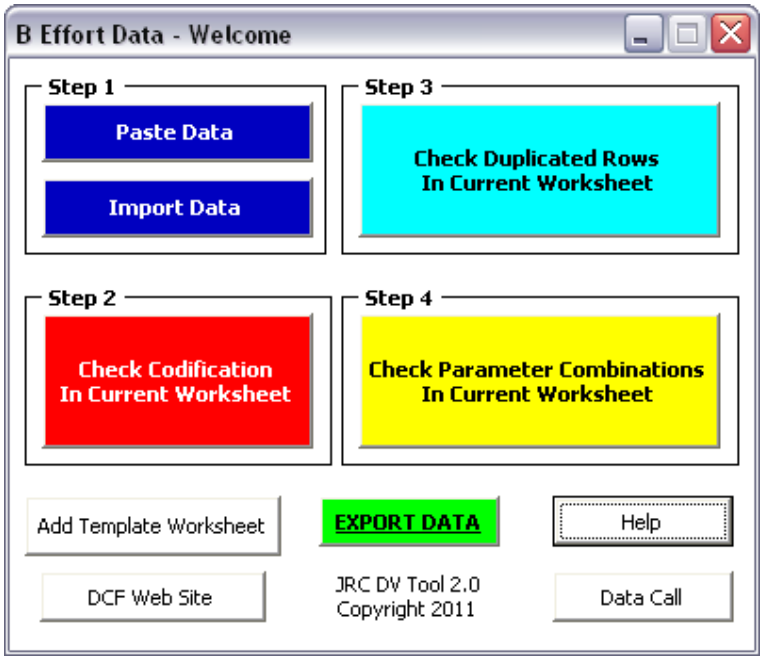



**Figure 16**: Welcome Screen of the DV tool for the Effort Data.

## 3.4. Step 3

For the B Effort table the aggregation level is defined by the following parameters:

COUNTRY – YEAR – QUARTER – VESSEL_LENGTH – GEAR – MESH_SIZE_RANGE – AREA – SPECON

### 3.5. Step 4

For **Step 4** please consult 2.5.

### 3.6. Messages

For **Messages** please consult 2.6.

## 4. C Spatial Effort Data

### 4.1. Welcome Screen

The Welcome Screen is the same as in the case of the Catch and Effort Data tables as shown in **Figure 17**. The window again consists of two main areas: a four step procedure, and a support area.

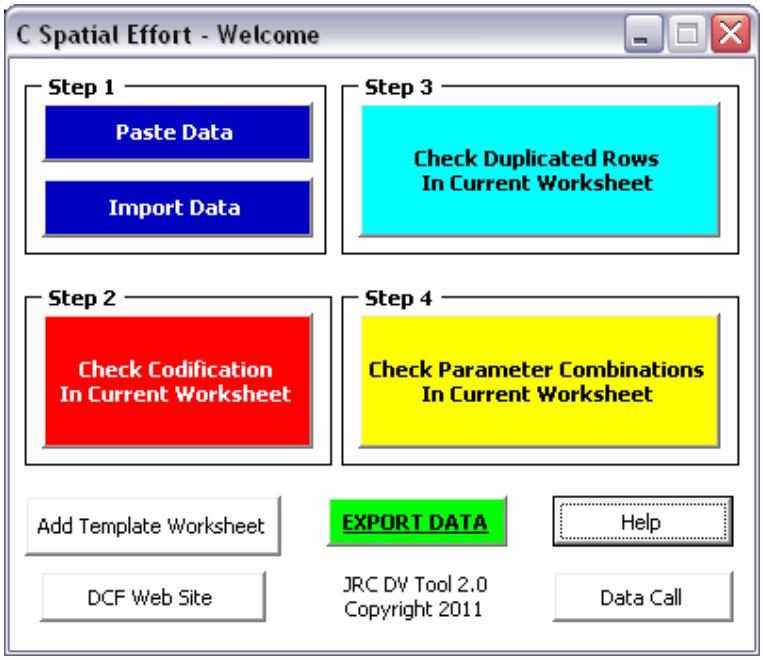



**Figure 17**: Welcome Screen of the DV tool for the Spatial Effort Data.

### 4.2. Step 1

For **Step 1**, please consider 2.2. The name of the default Worksheet is **EFF_03_SPEFF**.

### 4.3. Step 2

When clicking the red button, the *Check Codification* window appears and starts the procedure of checking if the data follows the codification scheme described in the effort data call.
The variables checked by this function are (case insensitive):

| | |
|---|---|
| COUNTRY | The valid codes are given in Appendix 1 of the data call. |
| YEAR | The entry should be a year, not after 2010 or earlier than 2003. |
| QUARTER | The quarter should be an integer of 1, 2, 3 or 4. |
| VESSEL_LENGTH | The valid codes are given in Appendix 2 of the data call. |
| GEAR | The valid codes are given in Appendix 3 of the data call. |

| | |
|---|---|
| MESH_SIZE_RANGE | The valid codes are given in Appendix 4 of the data call. |
| AREA | The valid codes are given in Appendix 5 of the data call. |
| SPECON | The valid codes are given in Appendix 6 of the data call. |
| RECTANGLE | The values should be ICES statistical rectangles (text 4 letters) for ICES Areas. No check is performed for CECAF Areas. |
| EFFECTIVE_EFFORT | The values should be numeric and not null or -1. |

### 4.4. Step 3

For the C Spatial Effort table the aggregation level is defined by the following parameters:

COUNTRY – YEAR – QUARTER – VESSEL_LENGTH – GEAR – MESH_SIZE_RANGE – AREA – SPECON - RECTANGLE

### 4.5. Step 4

For **Step 4** please consult 2.5.

### 4.6. Messages

For **Messages** please consult 2.6.

## 5. D Capacity Data

### 5.1. Welcome Screen

In the Welcome Screen of the Capacity data table there is a small difference comparing to the Catch, Effort Data and Spatial Effort data tables as shown in **Figure 18**. The window again consists of two main areas: a four step procedure (Check Combinations Step is not applicable), and a support area.

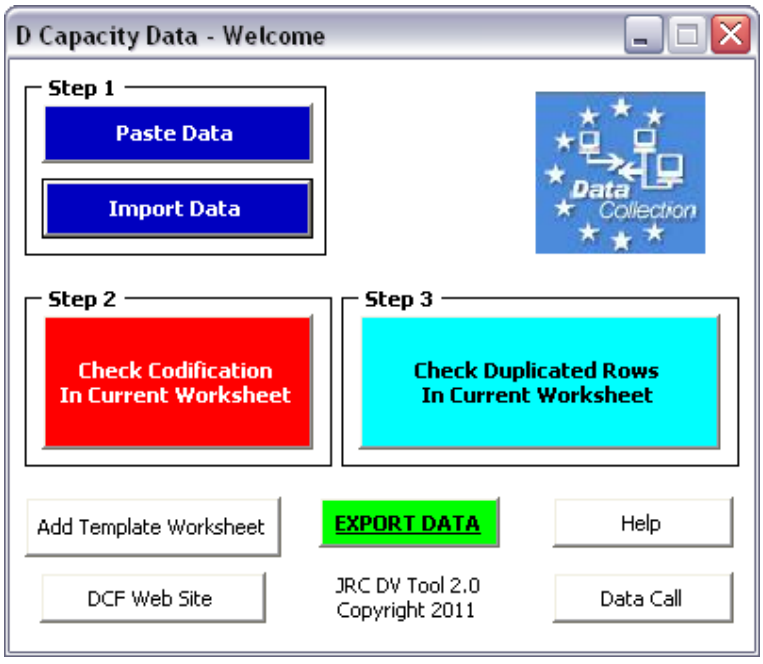



**Figure 18**: Welcome Screen of the DV tool for the Capacity Data.

### 5.2. Step 1

For **Step 1**, please consider 2.2. The name of the default Worksheet is **EFF_04_CAP**.

### 5.3. Step 2

When clicking the red button, the *Check Codification* window appears and starts the procedure of checking if the data follows the codification scheme described in the effort data call.
The variables checked by this function are (case insensitive):

| | |
|---|---|
| COUNTRY | The valid codes are given in Appendix 1 of the data call. |
| YEAR | The entry should be a year, not after 2010 or earlier than 2003. |
| VESSEL_LENGTH | The valid codes are given in Appendix 2 of the data call. |
| GEAR | The valid codes are REGGEAR or NONGEAR. |
| AREA | The valid codes are A, B or AB. |
| NO_VESSELS | The values should be numeric, not text. |
| FISHING_CAPACITY_KW | The values should be numeric, not text. |
| FISHING_CAPACITY_GT | The values should be numeric, not text. |

### 5.4. Step 3

For the D Capacity table the aggregation level is defined by the following parameters:

COUNTRY – YEAR – VESSEL_LENGTH – GEAR – AREA

### 5.5. Messages

For **Messages** please consult 2.6.

## 6. Correcting the Errors

Since the DV Tool is software designed for Excel, all the default functions of Excel are available. In order to correct the errors, the users can apply all the simple tools and functions provided by Excel.

There are three (3) colour definition for indicating errors or inconsistencies in the record fields:

Cells with red colour contain an invalid entry as a result of the **Step 2** procedure. These errors are critical because the can produce also errors in Step 4. The users should correct them prior proceeding to **Step 4**.

Cells with cyan colour are part of a set of duplicated records in the correspondence aggregation level. This indicates that the data are not aggregated according to the needs of the data call.

Cells with yellow colour are part of an invalid combination of parameter values. If the record contains an erroneous parameter combination, four (4) Cells will be coloured yellow, the **GEAR**, **MESH_SIZE_RANGE**, **AREA**, **SPECON** fields.

When a cell contains an invalid entry, apart from the colour notation, the whole record is pushed to the top of the Worksheet. This way all the erroneous records are easily accessed by the user. To inform about the position of the record in the original file just choose **Original Row Numbers.**

**Important:** After correcting the errors, **always** click on the **Re-evaluate** button to ensure that the changes are correct.

## 6.1. Step 2

When an error occurs in the **Step 2** procedure, the relevant cell is highlighted by red colour. Click on the red cell to show the **Error Correction Form (**see **Figure 19**).

The form allows the user to click on a drop down list with the valid codes for the erroneous field (**Figure 20**), and provides the user with two options.

**Correct Current** Corrects only the entry in the current Cell with the chosen option from the list.

**Correct All** Corrects all the equal erroneous entries in the current Column with the chosen option from the list.

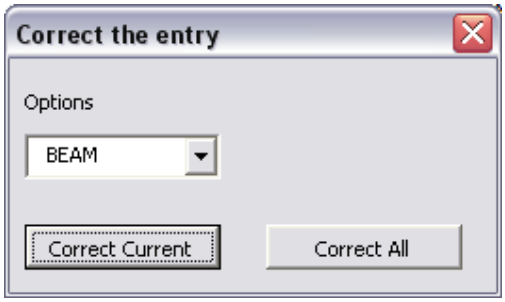



**Figure 19**: Error Correction Form for a wrong gear entry.

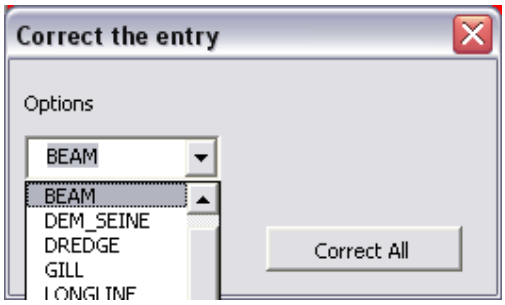



**Figure 20**: Available valid codes to choose for the correction of a wrong gear entry.

Another way to correct the errors is manually, by simply using **Find and Replace** or directly correcting the value of the cell. In case of doubts, the user can consult the **Valid Codes** in the View areas of the corresponding windows.

## 6.2. Step 3

An error in **Step 3** procedure is highlighted by cyan colour. In fact, all the parameters that form the aggregation level is coloured accordingly.

The user should manually identify the records that should be left in the data set and those that have to be deleted.

**Important:** Always repeat the checking procedure after correcting the errors using the **Re-evaluate** button.

### 6.3. Step 4

An error in **Step 4** procedure is highlighted by yellow colour. In fact, all the four parameters namely, **GEAR**, **MESH_SIZE_RANGE**, **AREA**, **SPECON**, will be coloured yellow.

In order to be informed about the valid combinations of these parameters, click on the **Valid Combinations** button. A Worksheet with all the available parameter combinations appears (**Figure 21**). The Worksheet has the **AutoFilter** option enabled. Click on the arrows next to the field names (Gear, Mesh_size_range, Specon, Area code) to filter the records according to the desired combination. For example, choosing **Gear** "LONGLINE", the available **Mesh_size_range** is only "-1".

After identifying the desired parameter combination, click on **Go Back to Data**, and return to the Worksheet with the data. Make all the necessary changes and proceed with the next records.

**Important:** Always repeat the checking procedure after correcting the errors using the **Re-evaluate** button. Only when the message given **Figure 9** appears the checking procedure has ended successfully.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | GEAR | MESH_SIZE_RANGE | SPECON | AREA |
| 2 | BEAM | 80-89 | -1 | 3as |
| 3 | BEAM | 80-89 | -1 | 3an |
| 4 | BEAM | 80-89 | -1 | 2 EU |
| 5 | BEAM | 80-89 | -1 | 4 |
| 6 | BEAM | 80-89 | -1 | 7d |
| 7 | BEAM | 80-89 | -1 | 7a |
| 8 | BEAM | 80-89 | -1 | 6a |
| 9 | BEAM | 90-99 | -1 | 3as |
| 10 | BEAM | 90-99 | -1 | 3an |
| 11 | BEAM | 90-99 | -1 | 2 EU |
| 12 | BEAM | 90-99 | -1 | 4 |
| 13 | BEAM | 90-99 | -1 | 7d |
| 14 | BEAM | 90-99 | -1 | 7a |
| 15 | BEAM | 90-99 | -1 | 6a |
| 16 | BEAM | 100-119 | -1 | 3as |
| 17 | BEAM | 100-119 | -1 | 3an |
| 18 | BEAM | 100-119 | -1 | 2 EU |
| 19 | BEAM | 100-119 | -1 | 4 |
| 20 | BEAM | 100-119 | -1 | 7d |

**Figure 21**: AutoFilter option to filter the valid combinations.

## 7. Export Data

In the Welcome Screen of the template Excel Workbooks, there is an option to **Export Data**. This button provide the option to save the data that are stored in the Workbook in new files, without the macros embedded, and with the option to split the files in order to reduce their size.

Clicking on this button, the user has two options **Figure 22**; the first is to split the Worksheets in smaller size files, and the second is to keep all the records in one file. When choosing yes, another window pops up (**Figure 23**) requiring the number of records per file. This option can help users that want to upload a large number of records (e.g. >40000 records) to the JRC's effort data bases.

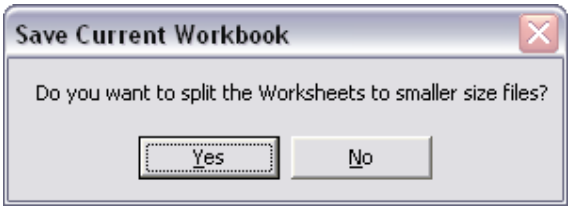



**Figure 22** Option to split the file to smaller size files.

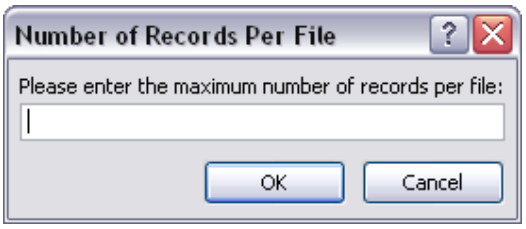



**Figure 23** Choose the number of records per file.

After completing the operation, a message appears (**Figure 23**) informing the users about the folder where the files are saved. By default, the application creates a new folder in the folder of the template file, with the name of the template and the date/time.

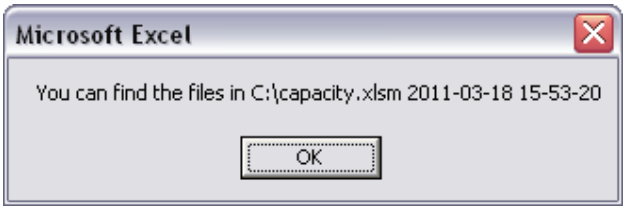



**Figure 24** Successful export message.

Each Worksheet is saved in a separate file, with the name of the Worksheet and no macro embedded. If the user chooses to split the worksheets in smaller files, the files are enumerated with the name of the Worksheet. The whole Worksheets are also saved as separate files. An example is given in **Figure 24**.

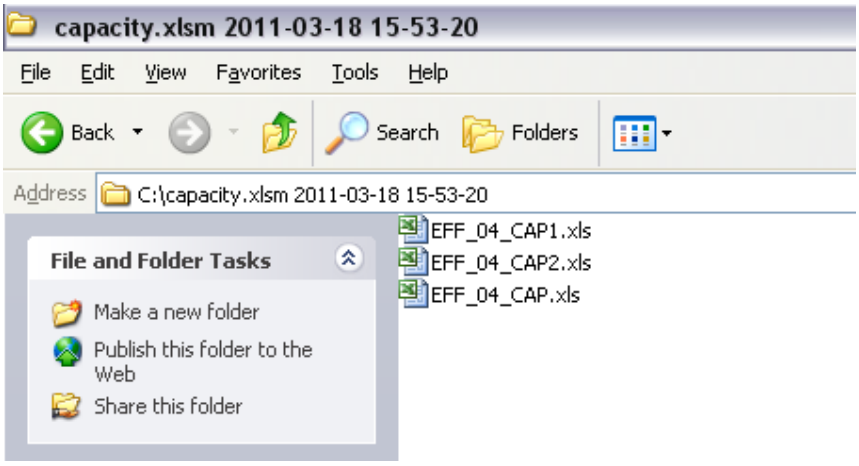



**Figure 25** Example of exported data in the created directory.

**Important:** Before exporting the data set always delete the Worksheets that do not contain checked data. Keep only the Worksheets that contain data that have been checked by the tool. In case you import the data, delete also the default empty Worksheet in order not to export this empty Worksheet.

Since the uploading facilities accept Office 2003 files and in case the number of records that the Worksheets contain are larger than 60000, you must split the data in smaller files with maximum 60000 records per file.

# 8. Office 2003

The current version of the DV Tool is also available for Office 2003 users. The four (4) Excel template files are named **catch.xls**, **effort.xls**, **speffort.xls** and **capacity.xls**.

Attention should be given to the restriction of a maximum number of 65535 rows (the first row is always the header). However, this is the THEORETICAL limit of number of rows per file in Excel 2003. In practise, and especially when dealing with large files, this number of rows can cause important computational time or even an Excel not responding status.

Users are strongly advised to use Excel 2007 when dealing with **more than 20000 rows** especially in A Catch data table, but also for the rest of the tables. The reasons are that the file size is getting large enough for common machines, the A Catch table has a large number of columns, the application will runner slower and the export data option could potentially not finish successfully. One other option is first to split manually the data set to smaller size and then insert or paste to the template files. Apply the DV tool separately for each one of these smaller data sets (e.g. per year data sets).

The environment is the same except for the option to recall and show the **Welcome Screen** from the Office Ribbon since this is not available for previous than Office 2007 versions. In order to re-show the Welcome Screen, simple use this key combination: **Ctrl + J** or just save and re-open the template.

To change the security settings for the Macros go to Tools → Macro → Security → Security Level → Medium.

# 9. Important Notes

- This document is a manual for the current version of the DV Tool 2.0.
- The users are strongly advised to download and install the latest version of Microsoft's Office Service Pack, which can be founded here.
- The DV Tool was tested successfully on Windows XP Service Pack 3 operating systems.
- The code of the application and the Worksheets containing the Valid Codes, and Valid Combinations are password protected.
- When dealing with large number of records please, be patient when checking or exporting data.
- When the Office 2003 versions are used, the number of total rows to check is limited to 65535. However, the Office 2007 version can handle over 1 million rows of data.
- For questions, suggestions, problems, bug reporting, support, send a direct email to the following email address: nikolaos.mitrakis@jrc.ec.europa.eu

# 10. Disclaimer of warranty

THE APPLICATION AND MANUAL ARE PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, WARRANTIES THAT THE CODE IS FREE OF DEFECTS, MERCHANTABLE,

FIT FOR A PARTICULAR PURPOSE OR NON-INFRINGING. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE CODE IS WITH YOU. SHOULD ANY CODE PROVE DEFECTIVE IN ANY RESPECT, YOU (NOT THE INITIAL DEVELOPER OR ANY OTHER CONTRIBUTOR) ASSUME THE COST OF ANY NECESSARY SERVICING, REPAIR OR CORRECTION. NO USE OF ANY CODE, APPLICATION OR MANUAL IS AUTHORIZED HEREUNDER EXCEPT UNDER THIS DISCLAIMER.

European Commission

**EUR 24845 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen**
Title: JRC Data Validation Tool for the Fisheries Effort Data Call through the EU Data Collection Regulation Framework (DCF) - DV Tool 2.0 - User Manual
Author: Nikolaos Mitrakis
Luxembourg: Publications Office of the European Union
2011 –21 pp. – 21 x 29.7 cm
EUR – Scientific and Technical Research series – ISSN 1831-9424 (online), ISSN 1018-5593 (print)
ISBN 978-92-79-20442-5
doi:10.2788/27561

**Abstract**
The Data Validation (DV) tool is a set of macros developed in Visual Basic for Applications (VBA) and embedded in specifically designed template Excel Workbooks for the effort data call. The main purpose of this tool is to facilitate and support the Member States in uploading data which meet the requirements defined by DG Mare in the official DCF data call for fishing effort regime evaluations by STECF (Council Regulation 199/2008). The use of these Excel Template files is not mandatory. However, the data validations checks performed by the DV tool can highly reduce the number of erroneous records the file to be uploaded to the DCF web site may contain, and hence facilitate the uploading procedure. The current version of the DV tool is 2.0. The purpose of this user manual is to provide the concept of this tool and guidelines for its use.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

JRC
EUROPEAN COMMISSION

LB-NA-24845-EN-N

ISBN 978-92-79-20442-5